

Homogenization of Dissipative, Noisy, Hamiltonian Dynamics

Jeremiah Birrell^a, Jan Wehr^{a,b}

^a*Department of Mathematics,*

^b*Program in Applied Mathematics*

University of Arizona

Tucson, AZ, 85721, USA

Abstract

We study the dynamics of a class of Hamiltonian systems with dissipation, coupled to noise, in a singular (small mass) limit. We derive the homogenized equation for the position degrees of freedom in the limit, including the presence of a *noise-induced drift* term. We prove convergence to the solution of the homogenized equation in an L^p -norm.

Keywords: Hamiltonian system, homogenization, small mass limit, noise-induced drift

2010 MSC: 60H10, 82C31

1. Introduction

In the simplest case, the motion of a diffusing particle of non-zero mass, m , is governed by a stochastic differential equation (SDE), of the form

$$dq_t = v_t dt, \quad m dv_t = -\gamma v_t dt + \sigma dW_t, \quad (1.1)$$

where γ and σ are the dissipation (or drag) and diffusion coefficients respectively and W_t is a Wiener process. The study of diffusive systems in the limit $m \rightarrow 0$ was initiated by Smoluchowski in [1] and continued by Kramers in [2]. The field has grown to explore a large array of models and phenomena, including coupled fluid-particle systems [3], relativistic diffusion [4, 5], and a variety of processes and convergence modes on manifolds [6, 7, 8, 9, 10, 11, 12, 13]. History of the subject and a review of the early literature can be found in [14].

Such problems can be classified under the broad umbrella of homogenization, for which [15] is an excellent reference.

Recently, there has been an increased interest in the phenomenon of *noise-induced drift*, which arises when the drag and noise coefficients are state dependent. In such cases, the equation governing the process in limit $m \rightarrow 0$ possesses an additional drift term that was not present in the original system. First derived in [16], this has been observed experimentally in [17] and derived rigorously for one dimensional systems [18], systems satisfying the fluctuation-dissipation relation [16], in Euclidean space of arbitrary dimension [19, 20], and on compact Riemannian manifolds of arbitrary dimension [21]. Further references to work on the phenomenon of noise-induced drift are found in [19].

Statistical mechanics of fluctuating systems, as reviewed in [22, 23], covers systems more general than those governed by the Hamiltonians with quadratic kinetic energy,

$$H(q, p) = \frac{\|p\|^2}{2m} + V(q), \quad (1.2)$$

but to this point, the study of noise-induced drift has been restricted to Hamiltonians quadratic in p . In this paper, we extend the theory to a large class of Hamiltonian systems generalizing Eq. (1.2). We prove that solutions to these more general Hamiltonian systems converge in an L^p -norm to solutions of a homogenized limiting equation with a noise-induced drift term, for which we derive an explicit formula. This is a far-reaching generalization of the classical question about the $m \rightarrow 0$ limit of the equation Eq. (1.1).

1.1. Dissipative Hamiltonian System with Noise

Here, we review the basic equations and properties of dissipative, noisy Hamiltonian systems. See also [22]. Given a time-dependent Hamiltonian $H(t, x)$ which is C^1 jointly in t and $x = (q, p) \in \mathbb{R}^n \times \mathbb{R}^n$, a positive-semidefinite continuous matrix-valued function $\Gamma(t, x)$, the matrix

$$\Pi = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad (1.3)$$

and a continuous vector field $G(t, x)$, we first consider the following deterministic equation.

$$\dot{x}_t = -\Gamma(t, x_t) \nabla H(t, x_t) + \Pi \nabla H(t, x_t) + G(t, x_t). \quad (1.4)$$

This equation describes the dynamics of a dissipative Hamiltonian system with drag matrix Γ and external forcing G . We will refer to q as the position degrees of freedom and to p as the momentum degrees of freedom.

The rate of change of the Hamiltonian along a solution is given by

$$\begin{aligned} & \frac{d}{dt}H(t, x_t) \\ &= \partial_t H(t, x_t) - \nabla H(t, x_t) \cdot \Gamma(t, x_t) \nabla H(t, x_t) \\ & \quad + \nabla H(t, x_t) \cdot \Pi \nabla H(t, x_t) + \nabla H(t, x_t) \cdot G(t, x_t) \\ & \leq \partial_t H(t, x_t) + \nabla H(t, x_t) \cdot G(t, x_t), \end{aligned} \tag{1.5}$$

where we used the anti-symmetry of Π and the positive semi-definiteness of Γ . In particular, if H is time independent and G vanishes then the energy is non-increasing and if Γ also vanishes then energy is conserved. This justifies the interpretation of Eq. (1.4) as a dissipative Hamiltonian system with external forcing G and drag matrix Γ .

We specialize to the case where the dissipation and external force enter only the momentum equation:

$$\Gamma(t, x) = \begin{pmatrix} 0 & 0 \\ 0 & \gamma(t, x) \end{pmatrix}, \tag{1.6}$$

and $G(t, x) = (0, F(t, x))$. With this, the dissipation couples linearly to the generalized velocity $v = \nabla_p H$, since the equations are now

$$\dot{q}_t = \nabla_p H(t, x_t), \quad \dot{p}_t = -\gamma(t, x_t) \nabla_p H(t, x_t) - \nabla_q H(t, x_t) + F(t, x_t). \tag{1.7}$$

We will be interested in Hamiltonians of the form

$$H(t, q, p) = K(t, q, p - \psi(t, q)) + V(t, q) \tag{1.8}$$

where $K = K(t, q, z)$ and $V = V(t, q)$ are C^2 , \mathbb{R} -valued functions, K is non-negative, and ψ is a C^2 , \mathbb{R}^n -valued function.

Remark 1. *The momentum-dependent term, K , and the momentum-independent term, V , into which H is split, do not have to carry with it the physical interpretation of kinetic and potential energy respectively, though we will use that terminology. The splitting will become constrained (though not quite unique) by further assumptions we will make below, but at this point it is largely arbitrary.*

We now define a family of scaled Hamiltonians, parameterized by $\epsilon > 0$:

$$H^\epsilon(t, q, p) \equiv K^\epsilon(t, q, p) + V(t, q) \equiv K(t, q, (p - \psi(t, q))/\sqrt{\epsilon}) + V(t, q). \quad (1.9)$$

Our motivation for the scaling used to define H^ϵ is the Hamiltonian of a classical particle coupled to a magnetic field, in which case

$$H(t, q, p) = \frac{\|p - e\phi(t, q)\|^2}{2m} + eV(t, q), \quad (1.10)$$

where e is the charge, ϕ is the vector potential, and V is the electrostatic potential. Our definition of H^ϵ is then equivalent to replacing m with ϵm . Hence, taking $\epsilon \rightarrow 0^+$ is equivalent to the small mass limit, $m \rightarrow 0$. In particular, for vanishing vector potential, this formalism will cover the small mass limit of a Newtonian particle in a potential.

Adding a noise term to the momentum components of Hamilton's equations, we arrive at the following family of SDEs:

$$dq_t^\epsilon = \nabla_p H^\epsilon(t, x_t^\epsilon) dt, \quad (1.11)$$

$$dp_t^\epsilon = (-\gamma(t, x_t^\epsilon) \nabla_p H^\epsilon(t, x_t^\epsilon) - \nabla_q H^\epsilon(t, x_t^\epsilon) + F(t, x_t^\epsilon)) dt + \sigma(t, x_t^\epsilon) dW_t, \quad (1.12)$$

where $\sigma : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times k}$ is continuous and W_t is a \mathbb{R}^k -valued Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual conditions [24]. In this paper we investigate the behavior of x_t^ϵ in the limit $\epsilon \rightarrow 0^+$ and derive a homogenized SDE satisfied by the limiting position process, q_t .

1.2. Summary of the Main Result

To prove our main theorem, we will require several assumptions — Assumptions 1-7 below. Not all of these are required for every result, so we state them only as needed. These assumptions will constrain the initial conditions, the analytical properties and form of the Hamiltonian, H , the drag matrix, γ , and the noise coefficients, σ . In particular, we will eventually require γ to be independent of p and its eigenvalues to satisfy a positive lower bound. The latter coercivity requirement will be crucial in proving the kinetic energy and momentum bounds in Section 2. Under these assumptions we will prove the following main result:

Let x_t^ϵ be a family of solutions to the SDE Eq. (1.11)-Eq. (1.12) with initial condition $x_0^\epsilon = (q_0^\epsilon, p_0^\epsilon)$. In this paper, we work under the assumptions that

a unique solution exists for all $t \geq 0$ (i.e. there are no explosions). See Appendix B for assumptions that guarantee this.

Let q_t be the solution to the SDE

$$dq_t = \tilde{\gamma}^{-1}(t, q_t)(-\partial_t \psi(t, q_t) - \nabla_q V(t, q_t) + F(t, q_t, \psi(t, q_t)))dt + S(t, q_t)dt + \tilde{\gamma}^{-1}(t, q_t)\sigma(t, q_t, \psi(t, q_t))dW_t \quad (1.13)$$

with initial condition q_0 , where $\tilde{\gamma}$ and $S(t, q)$ are defined in Eq. (3.3) and Eq. (3.26) respectively.

Suppose that for all $\epsilon > 0$ and all $p > 0$ we have $E[\|q_0^\epsilon\|^p] < \infty$, $E[\|q_0\|^p] < \infty$, and $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$. Then for any $T > 0$, $p > 0$, $0 < \beta < p/2$ we have

$$E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \text{ and } E\left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p\right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (1.14)$$

The drift term, $S(t, q)$, which appears in the limiting equation is called the *noise-induced drift* and is nonzero when σ is nonzero and a particular combination of quantities related to the kinetic energy, γ , and ψ have non-trivial state dependence. Other works studying the small mass limit of inertial systems, both in Euclidean space [18, 16, 19, 20], and on manifolds [21], have found analogous phenomena.

2. Kinetic Energy and Momentum Bounds

In this section, we derive bounds on the behavior of kinetic energy in the limit $\epsilon \rightarrow 0^+$ and prove a convergence result for the canonical momentum, the first formula in Eq. (1.14). We will require some assumptions on the structure of the Hamiltonian. As usual, here and in the sequel, generic symbols, denoting constants, such as C , M etc., do not have to have the same value in all equations.

Assumption 1. *We assume that the Hamiltonian has the form given in Eq. (1.8) where K and ψ are C^2 and K is non-negative. For every $T > 0$, we assume the following bounds hold on $[0, T] \times \mathbb{R}^{2n}$:*

1. *There exist $C > 0$ and $M > 0$ such that*

$$\max\{|\partial_t K(t, q, z)|, \|\nabla_q K(t, q, z)\|\} \leq M + CK(t, q, z). \quad (2.1)$$

2. There exist $c > 0$ and $M \geq 0$ such that

$$\|\nabla_z K(t, q, z)\|^2 + M \geq cK(t, q, z). \quad (2.2)$$

3. For every $\delta > 0$ there exists an $M > 0$ such that

$$\max \left\{ \|\nabla_z K(t, q, z)\|, \left(\sum_{ij} |\partial_{z_i} \partial_{z_j} K(t, q, z)|^2 \right)^{1/2} \right\} \leq M + \delta K(t, q, z). \quad (2.3)$$

The following lemma, which we state without proof, presents one general class of Hamiltonians that satisfy Assumption 1, but our main results do *not* assume that the Hamiltonian takes this form. In the statement of the Lemma, as well as in the remainder of the paper, we use the summation convention.

Lemma 2.1. *Suppose K has the following form:*

$$K(t, q, z) = \sum_{l=k_1}^{k_2} d_l(t, q) (A^{ij}(t, q) z_i z_j)^l \quad (2.4)$$

where $1 \leq k_1 \leq k_2$ are integers, $A = A^{ij}$ is a C^2 , positive-definite $n \times n$ -matrix-valued function, and d_l are C^2 and non-negative. Also, assume that for every $T > 0$, A and d_l satisfy the following bounds on $[0, T] \times \mathbb{R}^{2n}$:

1. There exists $\tilde{C} > 0$ such that $\|A\| \leq \tilde{C}$, $|d_l| \leq \tilde{C}$, $\|\partial_t A\| \leq \tilde{C}$, $|\partial_t d_l| \leq \tilde{C}$, $\|\partial_{q^i} A\| \leq \tilde{C}$, and $|\partial_{q^i} d_l| \leq \tilde{C}$ for all i, l .
2. There exists $\tilde{c} > 0$ such that $d_{k_1} \geq \tilde{c}$, $d_{k_2} \geq \tilde{c}$, and the eigenvalues of A are uniformly bounded below by \tilde{c} .

Then Eq. (2.1)-Eq. (2.3) from Assumption 1 hold.

We will also need the following assumptions concerning the potential, dissipation, noise, external forcing, and initial conditions:

Assumption 2. *For every $T > 0$, we assume that the following hold uniformly on $[0, T] \times \mathbb{R}^n$:*

1. V is C^2 and there exists $C > 0$ such that

$$\|\nabla_q V(t, q)\| \leq C. \quad (2.5)$$

2. The eigenvalues of γ are bounded below by some $\lambda > 0$.
3. γ , F , and σ are bounded.
4. There exists $C > 0$ such that the (random) initial conditions satisfy $K^\epsilon(0, x_0^\epsilon) \leq C$ for all $\epsilon > 0$ and all $\omega \in \Omega$.

We now state and prove the kinetic energy bound which underlies our main result.

Lemma 2.2. *Under Assumptions 1 and 2, for any $q \in \mathbb{N}$, $q \geq 1$ and any $T > 0$ there exist $\alpha_0 > 0$, $\epsilon_0 > 0$ and $\kappa > 0$ such that for all $0 < \alpha \leq \alpha_0$, $0 < \epsilon \leq \epsilon_0$, $0 \leq t \leq T$ we have the P -a.s. inequality*

$$K^\epsilon(t, x_t^\epsilon)^q \leq \frac{\kappa}{\alpha} + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \quad (2.6)$$

Proof. Take $T > 0$, $q \in \mathbb{N}$, $q \geq 1$, and apply Itô's formula to $K^\epsilon(t, x_t^\epsilon)^q$:

$$\begin{aligned} K^\epsilon(t, x_t^\epsilon)^q &= K^\epsilon(0, x_0^\epsilon)^q + q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_s K^\epsilon(s, x_s^\epsilon) ds \\ &+ q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_q K^\epsilon(s, x_s^\epsilon) \cdot \nabla_p H^\epsilon(s, x_s^\epsilon) ds \\ &- q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot \gamma(s, x_s^\epsilon) \nabla_p H^\epsilon(s, x_s^\epsilon) ds \\ &+ q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot (-\nabla_q H^\epsilon(s, x_s^\epsilon) + F(s, x_s^\epsilon)) ds \\ &+ q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \\ &+ \frac{q(q-1)}{2} \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-2} \partial_{p_i} K^\epsilon(s, x_s^\epsilon) \partial_{p_j} K^\epsilon(s, x_s^\epsilon) \sum_\rho \sigma_{i\rho}(s, x_s^\epsilon) \sigma_{j\rho}(s, x_s^\epsilon) ds \\ &+ \frac{q}{2} \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_{p_i} \partial_{p_j} K^\epsilon(s, x_s^\epsilon) \sum_\rho \sigma_{i\rho}(s, x_s^\epsilon) \sigma_{j\rho}(s, x_s^\epsilon) ds. \end{aligned} \quad (2.7)$$

Since $\nabla_p H^\epsilon = \nabla_p K^\epsilon$, the second line cancels with a part of the fourth, elimi-

nating the terms involving $\nabla_q K^\epsilon$ and resulting in

$$\begin{aligned}
K^\epsilon(t, x_t^\epsilon)^q &= K^\epsilon(0, x_0^\epsilon)^q + q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_s K^\epsilon(s, x_s^\epsilon) ds \\
&\quad - q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot \gamma(s, x_s^\epsilon) \nabla_p K^\epsilon(s, x_s^\epsilon) ds \\
&\quad + q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot (-\nabla_q V(s, x_s^\epsilon) + F(s, x_s^\epsilon)) ds \\
&\quad + \frac{q(q-1)}{2} \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-2} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot \Sigma(s, x_s^\epsilon) \nabla_p K^\epsilon(s, x_s^\epsilon) ds \\
&\quad + \frac{q}{2} \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_{p_i} \partial_{p_j} K^\epsilon(s, x_s^\epsilon) \Sigma_{ij}(s, x_s^\epsilon) ds \\
&\quad + q \int_0^t K^\epsilon(s, x_s^\epsilon)^{q-1} \nabla_p K^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s,
\end{aligned} \tag{2.8}$$

where we defined $\Sigma_{ij} = \sum_\rho \sigma_{i\rho} \sigma_{j\rho}$, i.e. $\Sigma = \sigma \sigma^T$.

Since the function K is evaluated at $z = (p - \psi(t, q))/\sqrt{\epsilon}$, will use the abbreviated notation

$$(\partial_{z^i} K)^\epsilon(t, x) = \partial_{z^i} K(t, q, (p - \psi(t, q))/\sqrt{\epsilon}) \tag{2.9}$$

and similarly for $(\nabla_z K)^\epsilon$, $(\partial_{z^i} \partial_{z^j} K)^\epsilon$, etc. Using Itô's formula again, along

with Assumption 2, for any $\alpha > 0$ and any $t \in [0, T]$ we have

$$e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q = K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \quad (2.10)$$

$$\begin{aligned} & + q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_s K^\epsilon(s, x_s^\epsilon) ds \\ & - \frac{q}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \gamma(s, x_s^\epsilon) (\nabla_z K)^\epsilon(s, x_s^\epsilon) ds \\ & + \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot (-\nabla_q V(s, x_s^\epsilon) + F(s, x_s^\epsilon)) ds \\ & + \frac{q(q-1)}{2\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \Sigma(s, x_s^\epsilon) (\nabla_z K)^\epsilon(s, x_s^\epsilon) ds \\ & + \frac{q}{2\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\partial_{z_i} \partial_{z_j} K)^\epsilon(s, x_s^\epsilon) \Sigma_{ij}(s, x_s^\epsilon) ds \\ & + \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \\ & \leq K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \end{aligned} \quad (2.11)$$

$$\begin{aligned} & + q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \partial_s K^\epsilon(s, x_s^\epsilon) ds \\ & - \frac{q\lambda}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 ds \\ & + \frac{q}{\sqrt{\epsilon}} \|\nabla_q V + F\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\| ds \\ & + \frac{q(q-1)}{2\epsilon} \|\Sigma\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 ds \\ & + \frac{q}{2\epsilon} \|\Sigma\|_{F,\infty} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \left(\sum_{ij} (\partial_{z_i} \partial_{z_j} K)^\epsilon(s, x_s^\epsilon)^2 \right)^{1/2} ds \\ & + \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \end{aligned}$$

Here and in the following we will use $\|Y\|_F$ to denote the Frobenius (or Hilbert-Schmidt) norm of a matrix Y , i.e. $\|Y\| = (\sum_{ij} Y_{ij}^2)^{\frac{1}{2}}$. We write $\|Y\|_\infty \equiv \sup_{(t,q) \in [0,T] \times \mathbb{R}^n} \|Y(t,q)\|$ for any matrix or vector valued quantity Y and similarly for $\|\cdot\|_{F,\infty}$. The implied value of T will be clear from the context.

Now, Assumption 1 implies that for any $\delta > 0$, there exist $C > 0$, $c > 0$, and $M > 0$ such that

$$\begin{aligned}
& e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q \tag{2.12} \\
& \leq K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
& \quad + q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + CK^\epsilon(s, x_s^\epsilon)) ds \\
& \quad - \frac{q\lambda}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (cK^\epsilon(s, x_s^\epsilon) - M) ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} \|\nabla_q V + F\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + \delta K^\epsilon(s, x_s^\epsilon)) ds \\
& \quad + \frac{q(q-1)}{2\epsilon} \|\Sigma\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} (M + \delta K^\epsilon(s, x_s^\epsilon))^2 ds \\
& \quad + \frac{q}{2\epsilon} \|\Sigma\|_{F,\infty} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + \delta K^\epsilon(s, x_s^\epsilon)) ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s.
\end{aligned}$$

In the estimate that follows, the first two terms and the last term of the above expression will be left unchanged. To estimate the remaining terms, we will use the elementary inequalities:

$$K^{q-1} \leq \left(\frac{M}{\delta}\right)^{q-1} + \frac{\delta}{M} K^q \tag{2.13}$$

$$K^{q-2} \leq \left(\frac{M}{\delta}\right)^{q-2} + \left(\frac{\delta}{M}\right)^2 K^q. \tag{2.14}$$

The first inequality holds for every $q \geq 1$ and the second for every $q \geq 2$. Note that for $q = 1$, the term containing $K^\epsilon(s, x_s^\epsilon)^{q-2}$ vanishes. Applying these inequalities, we obtain

$$\begin{aligned}
& K^\epsilon(t, x_t^\epsilon)^q \tag{2.15} \\
& \leq e^{-\alpha t/\epsilon} K^\epsilon(0, x_0^\epsilon)^q + \frac{D}{\alpha} - \frac{d}{\epsilon} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s,
\end{aligned}$$

where

$$D = qM \left(\frac{M}{\delta} \right)^{q-1} \left[\lambda + \epsilon + \sqrt{\epsilon} \|\nabla_q V + F\|_\infty + \frac{1}{2} \|\Sigma\|_{F,\infty} \right] + q(q-1)M^2(M/\delta)^{q-2} \|\Sigma\|_\infty, \quad (2.16)$$

$$d = qc\lambda - \alpha - qC\epsilon - q\delta\sqrt{\epsilon} \|\nabla_q V + F\|_\infty - \frac{q\delta}{2} \|\Sigma\|_{F,\infty} - q(q-1)\delta^2 \|\Sigma\|_\infty - q(q-1)\delta^2 \|\Sigma\|_\infty - \left(q\lambda + q\epsilon + q\sqrt{\epsilon} \|\nabla_q V + F\|_\infty + \frac{q}{2} \|\Sigma\|_{F,\infty} \right) \delta. \quad (2.17)$$

For all ϵ, δ, α sufficiently small, d is non-negative, and hence

$$K^\epsilon(t, x_t^\epsilon)^q \leq K^\epsilon(0, x_0^\epsilon)^q + \frac{D}{\alpha} + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \quad (2.18)$$

By Assumption 2, $K^\epsilon(0, x_0^\epsilon)$ is bounded, so we are done. \square

We will use this bound to prove several results about the behavior of the kinetic energy and momentum as $\epsilon \rightarrow 0^+$.

2.1. Integrability of the Kinetic Energy

Lemma 2.3. *Under Assumptions 1 and 2, $E[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p]$ is finite for any $T > 0$, $\epsilon > 0$, and $p > 0$.*

Proof. Fix $T > 0$, $\epsilon > 0$. First, let $p > 2$. Given $M > 0$, define the stopping time $\tau_M = \inf\{t : K(t, x_t^\epsilon) = M\}$. By Assumption 2 we can take M large enough so that $K^\epsilon(0, x_0^\epsilon) < M$.

Let $t \leq T$ and raise Eq. (2.15) (with $q = 1$) to the p th power to obtain

$$\begin{aligned} & K^\epsilon(t \wedge \tau_M, (x^\epsilon)_t^{\tau_M})^p \\ & \leq C_1 + C_2 \left(\int_0^t 1_{s \leq \tau_M} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) ds \right)^p \\ & \quad + C_3 \left| \int_0^t 1_{s \leq \tau_M} e^{\alpha(s \wedge \tau_M)/\epsilon} (\nabla_z K)^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) \cdot \sigma(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) dW_s \right|^p \end{aligned} \quad (2.19)$$

where C_i are constants (that depend on ϵ and T).

Therefore, applying Hölder's inequality to the second term and the Burkholder-Davis-Gundy inequality to the third term, (see, for example, Theorem 3.28 in [24]), we obtain

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E[K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p] ds \end{aligned} \quad (2.20)$$

$$\begin{aligned} & + C_4 E \left[\left(\int_0^t 1_{r \leq \tau_M} e^{2\alpha(r \wedge \tau_M)/\epsilon} \|(\nabla_z K)^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})\|^2 \right. \right. \\ & \quad \left. \left. \times \|\sigma(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})\|^2 dr \right)^{p/2} \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds \\ & + C_5 E \left[\left(\int_0^t (1 + K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M}))^2 dr \right)^{p/2} \right]. \end{aligned} \quad (2.21)$$

By assumption, $p > 2$, so we can use Hölder's inequality again to obtain

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} K(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds \end{aligned} \quad (2.22)$$

$$\begin{aligned} & + C_5 T^{p/2-1} E \left[\int_0^t (1 + K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M}))^p dr \right] \\ & \leq C_6 + C_7 \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds \end{aligned} \quad (2.23)$$

where C_i are independent of t and M .

By the definition of τ_M ,

$$\sup_{s \in [0, t]} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \leq M \quad (2.24)$$

for all t . Therefore the integral in Eq. (2.23) is finite for all $t \leq T$ and Gronwall's inequality gives

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t \wedge \tau_M, (x^\epsilon)_t^{\tau_M})^p \right] \leq C_6 e^{C_7 T}. \quad (2.25)$$

The C_i are independent of M , so taking $M \rightarrow \infty$ and using the Monotone Convergence Theorem implies

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] \leq C_6 e^{C_7 T} < \infty. \quad (2.26)$$

This gives the result for $p > 2$. It follows for all $p > 0$ by an application of Hölder's inequality. \square

2.2. Supremum of the Expectation of the Kinetic Energy

Combining Lemmas 2.2 and 2.3 we can prove the following bound for the supremum of the expected value of the kinetic energy.

Proposition 2.1. *Under the Assumptions 1 and 2, for any $T > 0$, $q > 0$ we have*

$$\sup_{t \in [0, T]} E[K^\epsilon(t, x_t^\epsilon)^q] = O(1) \text{ as } \epsilon \rightarrow 0^+. \quad (2.27)$$

Proof. First take $T > 0$, $q \in \mathbb{N}$, $q \geq 1$. The following computation shows that

$$M_t \equiv \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \quad (2.28)$$

is a martingale (see [24]):

$$\begin{aligned} & E \left[\int_0^t \|e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 ds \right] \\ & \leq e^{2\alpha t/\epsilon} \|\sigma\|_\infty^2 t E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2(q-1)} (M + K^\epsilon(s, x_s^\epsilon))^2 \right] \\ & \leq 2e^{2\alpha t/\epsilon} \|\sigma\|_\infty^2 t \left(M^2 E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2(q-1)} \right] + E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2q} \right] \right) < \infty, \end{aligned} \quad (2.29)$$

where we used Assumption 1 and Lemma 2.3.

Therefore, taking the expectation of Eq. (2.6), we see that there exists $\kappa > 0$ such that for all $t \leq T$ and all α and ϵ sufficiently small, we have

$$E[K^\epsilon(t, x_t^\epsilon)^q] \leq \frac{\kappa}{\alpha}. \quad (2.30)$$

This proves the result for q a positive integer. The result then follows for arbitrary $q > 0$ by an application of Hölder's inequality. \square

Corollary 2.1. *We note that if the constants involved in the bounds from Assumptions 1-2 are valid uniformly for $t \in [0, \infty)$ (and not just $t \in [0, T]$) then we obtain the stronger bound*

$$\sup_{t \in [0, \infty)} E[K^\epsilon(t, x_t^\epsilon)^q] = O(1) \text{ as } \epsilon \rightarrow 0^+ \quad (2.31)$$

for any $q > 0$.

2.3. Expectation of the Supremum of the Kinetic Energy

We now have the ingredients to derive a bound on the expectation of the supremum of the kinetic energy. For this, we need to recall a special case of the Lemma 5.1 from [21]:

Lemma 2.4. *Let $V \in L^2_{loc}(dt)$ be \mathbb{R}^k -valued. For any $\alpha > 0$, $T \geq \delta > 0$ we have the P -a.s. bound*

$$\begin{aligned} & \sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)} V_s dW_s \right| \\ & \leq 5 \left(e^{-\alpha\delta} \sup_{t \in [0, T]} \left| \int_0^t V_r dW_r \right| + \max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t V_r dW_r \right| \right) \end{aligned} \quad (2.32)$$

where $N = \max\{k \in \mathbb{Z} : k\delta < T\}$.

Proposition 2.2. *Under Assumptions 1 and 2, for any $T > 0$, $p > 0$, $\beta > 0$ we have*

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] = O(\epsilon^{-\beta}) \text{ as } \epsilon \rightarrow 0^+. \quad (2.33)$$

Proof. By Lemma 2.2 with $q = 1$, there exist $\alpha > 0$ and $\kappa > 0$ such that for all ϵ sufficiently small and all $t \in [0, T]$, the following bound holds a.s.:

$$K^\epsilon(t, x_t^\epsilon) \leq \frac{\kappa}{\alpha} + \frac{1}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \quad (2.34)$$

We will first prove the proposition under the additional assumption $p > 2$. The general case $p > 0$ will follow by an application of Hölder's inequality.

$$\begin{aligned} & E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] \\ & \leq 2^{p-1} (\kappa/\alpha)^p + \frac{2^{p-1}}{\epsilon^{p/2}} E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right]. \end{aligned} \quad (2.35)$$

For any $T \geq \delta > 0$, Lemma 2.4 (with α/ϵ in place of α) implies

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \\
& \leq 5^p 2^{p-1} \left(e^{-p\alpha\delta/\epsilon} E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \right. \\
& \quad \left. + E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \right)
\end{aligned} \tag{2.36}$$

where $N = \max\{k \in \mathbb{Z} : k\delta < T\}$.

The Burkholder-Davis-Gundy inequality, applied to the first term on the right side of the inequality, implies existence of a constant $\tilde{C} > 0$ such that

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \\
& \leq \tilde{C} E \left[\left(\int_0^T \|(\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 dr \right)^{p/2} \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p E \left[\left(\int_0^T \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 dr \right)^{p/2} \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p E \left[\left(\int_0^T (M + K^\epsilon(s, x_s^\epsilon))^2 dr \right)^{p/2} \right].
\end{aligned} \tag{2.37}$$

In the last line, we used Assumption 1.

We have assumed $p > 2$, so we can use Hölder's inequality with exponents $p/(p-2)$ and $p/2$, to get

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p T^{p/2-1} E \left[\int_0^T (M + K^\epsilon(s, x_s^\epsilon))^p dr \right] \\
& \leq 2^{p-1} \tilde{C} \|\sigma\|_\infty^p T^{p/2} \left(M^p + \sup_{s \in [0, T]} E[K^\epsilon(s, x_s^\epsilon)^p] \right) \\
& = O(1)
\end{aligned} \tag{2.38}$$

as $\epsilon \rightarrow 0^+$ by Proposition 2.1.

We now work on the second term in Eq. (2.36). Using the fact that the ℓ^∞ -norm on \mathbb{R}^N is bounded by the $\ell^{\tilde{p}}$ norm for any $\tilde{p} \geq 1$, and then applying Hölder's inequality and the Burkholder-Davis-Gundy inequality, we derive the bound

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (2.39) \\
& \leq E \left[\left(\sum_{k=0}^{N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^{p\tilde{p}} \right)^{1/\tilde{p}} \right] \\
& \leq \left(\sum_{k=0}^{N-1} E \left[\sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^{p\tilde{p}} \right] \right)^{1/\tilde{p}} \\
& \leq \left(\sum_{k=0}^{N-1} \tilde{C} E \left[\left(\int_{k\delta}^{(k+2)\delta} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 dr \right)^{p\tilde{p}/2} \right] \right)^{1/\tilde{p}} \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left(\sum_{k=0}^{N-1} E \left[\left(\int_{k\delta}^{(k+2)\delta} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 dr \right)^{p\tilde{p}/2} \right] \right)^{1/\tilde{p}}.
\end{aligned}$$

Note that for $0 \leq k < N$ we have $0 \leq (k+2)\delta \leq (N+1)\delta \leq 2T$. So here, the time interval corresponding to $\|\cdot\|_\infty$, etc., can be taken to be $[0, 2T]$.

By assumption, $p\tilde{p} > 2$, so using Hölder's inequality again with exponents $p\tilde{p}/(p\tilde{p}-2)$ and $p\tilde{p}/2$, along with Assumption 1, we get

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (2.40) \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left(\sum_{k=0}^{N-1} (2\delta)^{p\tilde{p}/2-1} \int_{k\delta}^{(k+2)\delta} E[\|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^{p\tilde{p}}] dr \right)^{1/\tilde{p}} \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left((2\delta)^{p\tilde{p}/2} N \right)^{1/\tilde{p}} \sup_{s \in [0, (N+1)\delta]} E[(M + K^\epsilon(s, x_s^\epsilon))^{p\tilde{p}}]^{1/\tilde{p}}.
\end{aligned}$$

$N < T/\delta$, therefore

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (2.41) \\
& \leq 2^{p/2} \tilde{C}^{1/\tilde{p}} T^{1/\tilde{p}} \|\sigma\|_\infty^p \delta^{p/2-1/\tilde{p}} \sup_{s \in [0, 2T]} E[(M + K^\epsilon(s, x_s^\epsilon))^{p\tilde{p}}]^{1/\tilde{p}} \\
& = \delta^{p/2-1/\tilde{p}} O(1),
\end{aligned}$$

where we used Proposition 2.1.

Combining these results we see that for $\epsilon > 0$ sufficiently small and any $T \geq \delta > 0$, $\tilde{p} \geq 1$ we have

$$E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \leq e^{-p\alpha\delta/\epsilon} O(1) + \delta^{p/2-1/\tilde{p}} O(1) \quad (2.42)$$

where the big-O terms do not depend on δ .

Now let $0 < \xi < 1$ and choose $\delta = \epsilon^{1-\xi}$. Then

$$\begin{aligned} E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] &\leq 2^{p-1} (\kappa/\alpha)^p + \frac{2^{p-1}}{\epsilon^{p/2}} (e^{-p\alpha/\epsilon^\xi} O(1) + \epsilon^{(1-\xi)(p/2-1/\tilde{p})} O(1)) \\ &= 2^{p-1} (\kappa/\alpha)^p + \epsilon^{-p/2} e^{-p\alpha/\epsilon^\xi} O(1) + \epsilon^{(1-\xi)(p/2-1/\tilde{p})-p/2} O(1). \end{aligned} \quad (2.43)$$

For any $\beta > 0$ there exists $\tilde{p} \geq 1$ and $0 < \xi < 1$ such that

$$(1-\xi)(p/2-1/\tilde{p})-p/2 = -\xi \frac{p}{2} - \frac{1-\xi}{\tilde{p}} > -\beta. \quad (2.44)$$

so that the term $\epsilon^{(1-\xi)(p/2-1/\tilde{p})-p/2} O(1)$ diverges more slowly than $\epsilon^{-\beta}$. Also, $\epsilon^{-p/2} e^{-p\alpha/\epsilon^\xi} = o(1)$ for all $\xi > 0$. This proves the result for $p > 2$. The result for all $p > 0$ again follows by an application of Hölder's inequality. \square

2.4. Decay of Momentum

Starting in this section, we will assume that the difference between the canonical momentum and ψ is bounded by the kinetic energy in the following sense:

Assumption 3. *We assume that for every $T > 0$ there exists $c > 0$, $\eta > 0$ such that*

$$K(t, q, z) \geq c \|z\|^{2\eta} \quad (2.45)$$

on $[0, T] \times \mathbb{R}^{2n}$.

With the addition of Assumption 3, the bounds on the kinetic energy from Propositions 2.1 and 2.2 yield the following decay rates for the momentum to the submanifold defined by $p = \psi(t, q)$:

Lemma 2.5. *Under Assumptions 1-3, for any $T > 0$, $p > 0$ we have*

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+ \quad (2.46)$$

and for any $p > 0$, $T > 0$, $0 < \beta < p/2$ we have

$$E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (2.47)$$

Our final momentum decay rate result concerns a class of integrals with respect to products of the components of $u_t^\epsilon \equiv p_t^\epsilon - \psi(t, q_t^\epsilon)$.

Proposition 2.3. *Let $f : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^1 function, such that for every $T > 0$, f , $\partial_t f$, and $\nabla_q f$ are bounded on $[0, T] \times \mathbb{R}^n$. Define $u_t^\epsilon = p_t^\epsilon - \psi(t, q_t^\epsilon)$. Under Assumptions 1-3, for any $p > 0$, $T > 0$, $i, j = 1, \dots, n$ we have*

$$E\left[\sup_{t \in [0, T]} \left|\int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j)\right|^p\right] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+. \quad (2.48)$$

Proof. $f(s, q_s^\epsilon)$ is a C^1 -semimartingale. Therefore integration by parts gives

$$\begin{aligned} \int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j) &= f(t, q_t^\epsilon) (u_t^\epsilon)_i (u_t^\epsilon)_j - f(0, q_0^\epsilon) (z_0^\epsilon)_i (z_0^\epsilon)_j \\ &\quad - \int_0^t (u_s^\epsilon)_i (u_s^\epsilon)_j (\partial_s f(s, q_s^\epsilon) + \nabla_q f(s, q_s^\epsilon) \cdot \nabla_p H^\epsilon(s, x_s^\epsilon)) ds. \end{aligned} \quad (2.49)$$

Hence, for $p \geq 1$, using Assumption 3 we obtain:

$$\begin{aligned} &E\left[\sup_{t \in [0, T]} \left|\int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j)\right|^p\right] \\ &\leq 3^{p-1} \left(E\left[\sup_{t \in [0, T]} |f(t, q_t^\epsilon) (u_t^\epsilon)_i (u_t^\epsilon)_j|^p\right] + E[|f(0, q_0^\epsilon) (z_0^\epsilon)_i (z_0^\epsilon)_j|^p] \right. \\ &\quad \left. + E\left[\sup_{t \in [0, T]} \left|\int_0^t (u_s^\epsilon)_i (u_s^\epsilon)_j (\partial_s f(s, q_s^\epsilon) + \nabla_q f(s, q_s^\epsilon) \cdot \nabla_p K^\epsilon(s, x_s^\epsilon)) ds\right|^p\right] \right) \\ &\leq 3^{p-1} \left(2\|f\|_\infty^p E\left[\sup_{t \in [0, T]} \epsilon^p (K^\epsilon(t, x_t^\epsilon)/c)^{p/\eta}\right] \right. \\ &\quad \left. + E\left[\left(\int_0^T \|u_s^\epsilon\|^2 (\|\partial_s f\|_\infty + \|\nabla_q f\|_\infty \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)/\sqrt{\epsilon}\|) ds\right)^p\right] \right). \end{aligned} \quad (2.50)$$

Now using Assumption 1 and Proposition 2.2, for any $\beta > 0$ we find

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t f(s, q_s^\epsilon) d((p_s^\epsilon)^i (p_s^\epsilon)^j) \right|^p \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{p-1} E \left[\left(\int_0^T \epsilon (K^\epsilon(s, x_s^\epsilon)/c)^{1/\eta} \right. \right. \\
& \quad \left. \left. \times (\|\partial_s f\|_\infty + \epsilon^{-1/2} \|\nabla_q f\|_\infty (M + K^\epsilon(s, x_s^\epsilon))) ds \right)^p \right].
\end{aligned} \tag{2.51}$$

Hölder's inequality and Proposition 2.1 allow us to bound the second term:

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j) \right|^p \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{p-1} c^{-p/\eta} T^{p-1} \epsilon^p E \left[\int_0^T K^\epsilon(s, x_s^\epsilon)^{p/\eta} (\|\partial_s f\|_\infty \right. \\
& \quad \left. + \epsilon^{-1/2} \|\nabla_q f\|_\infty (M + K^\epsilon(s, x_s^\epsilon)))^p ds \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{2(p-1)} c^{-p/\eta} T^p \epsilon^p \sup_{t \in [0, T]} E [K^\epsilon(s, x_s^\epsilon)^{p/\eta} (\|\partial_s f\|_\infty^p \\
& \quad + \epsilon^{-p/2} \|\nabla_q f\|_\infty^p (M^p + K^\epsilon(s, x_s^\epsilon)^p))] \\
& = O(\epsilon^{p-\beta}) + 3^{2(p-1)} c^{-p/\eta} T^p \epsilon^p O(\epsilon^{-p/2}).
\end{aligned} \tag{2.52}$$

Taking $\beta = p/2$ gives the result when $p \geq 1$. The result for any $p > 0$ follows from Hölder's inequality. \square

3. Derivation of the Limiting Equation

In this section, we derive the equation satisfied by q_t^ϵ in the limit $\epsilon \rightarrow 0^+$. The actual convergence proof will be given in the following section. The derivation is an adaptation of the methods used in [19, 21]. We will need the following:

Assumption 4. *We assume that γ is C^1 and is independent of p .*

The starting point for the derivation is a rewriting of Hamilton's equation of motion in terms of the variables $u_t^\epsilon \equiv p_t^\epsilon - \psi(t, q_t^\epsilon)$:

$$d(u_t^\epsilon)_i = -\gamma_{ij}(t, q_t^\epsilon) \partial_{p_j} H^\epsilon(t, x_t^\epsilon) dt + (-\partial_{q^i} H^\epsilon(t, x_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \quad (3.1)$$

$$\begin{aligned} & -\partial_t \psi_i(t, q_t^\epsilon) dt - \partial_{q^k} \psi_i(t, q_t^\epsilon) \partial_{p_k} H^\epsilon(t, x_t^\epsilon) dt + \sigma_{ij}(t, x_t^\epsilon) dW_t^j \\ & = -\tilde{\gamma}_{ik}(t, q_t^\epsilon) \partial_{p_k} K^\epsilon(t, x_t^\epsilon) dt - (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ & \quad + (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt + \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho \end{aligned} \quad (3.2)$$

where

$$\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q). \quad (3.3)$$

The second and third terms in $\tilde{\gamma}$ together form an antisymmetric matrix, hence the eigenvalue bound for γ from Assumption 2 implies invertibility of $\tilde{\gamma}$. See Lemma A1.

This lets us solve for $\nabla_p H^\epsilon(t, x_t^\epsilon) dt$ to get

$$\begin{aligned} d(q_t^\epsilon)^i &= \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt \\ &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ & \quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho \\ & \quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) d(u_t^\epsilon)_j. \end{aligned} \quad (3.4)$$

$\tilde{\gamma}^{-1}(t, q_t^\epsilon)$ is pathwise C^1 , so integrating the last term by parts results in

$$\begin{aligned} -(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) d(u_t^\epsilon)_j &= -d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\ & \quad + (u_t^\epsilon)_j \partial_{q^l} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{p_l} H^\epsilon(t, x_t^\epsilon) dt. \end{aligned} \quad (3.5)$$

Therefore

$$\begin{aligned} d(q_t^\epsilon)^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ & \quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho \\ & \quad - d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\ & \quad + (u_t^\epsilon)_j \partial_{q^l} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{p_l} H^\epsilon(t, x_t^\epsilon) dt. \end{aligned} \quad (3.6)$$

In order to homogenize $(u_t^\epsilon)_j \partial_{p_l} H^\epsilon(t, x_t^\epsilon) dt$, we make the additional assumption:

Assumption 5. We assume that K has the form

$$K(t, q, z) = \tilde{K}(t, q, A^{ij}(t, q)z_i z_j) \quad (3.7)$$

where $\tilde{K}(t, q, \zeta)$ is C^2 and non-negative on $[0, \infty) \times \mathbb{R}^n \times [0, \infty)$ and $A(t, q)$ is a C^2 function whose values are symmetric $n \times n$ -matrices. We also assume that for every $T > 0$, the eigenvalues of A are bounded above and below by some constants $C > 0$ and $c > 0$ respectively, uniformly on $[0, T] \times \mathbb{R}^n$.

We will write \tilde{K}' for $\partial_\zeta \tilde{K}$ and will use the abbreviation $\|z\|_A^2$ for $A^{ij}(t, q)z_i z_j$ when the implied values of t and q are apparent from the context.

With this assumption,

$$\begin{aligned} (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) &= \partial_{q^i} \tilde{K}(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) \\ &\quad + \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) \partial_{q^i} A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_k (u_t^\epsilon)_l / \epsilon \end{aligned} \quad (3.8)$$

and

$$\partial_{p_l} H^\epsilon(t, x_t^\epsilon) = \frac{2}{\epsilon} A^{lk}(t, q_t^\epsilon) \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) (u_t^\epsilon)_k. \quad (3.9)$$

To simplify $(u_t^\epsilon)_j \partial_{p_l} H^\epsilon(t, x_t^\epsilon) dt$, we compute

$$\begin{aligned} d((u_t^\epsilon)_i (u_t^\epsilon)_j) &= (u_t^\epsilon)_i d(u_t^\epsilon)_j + (u_t^\epsilon)_j d(u_t^\epsilon)_i + d[u_i^\epsilon, u_j^\epsilon]_t \\ &= (- (u_t^\epsilon)_i \tilde{\gamma}_{jk}(t, q_t^\epsilon) - (u_t^\epsilon)_j \tilde{\gamma}_{ik}(t, q_t^\epsilon)) \frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l dt \\ &\quad - (u_t^\epsilon)_i (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt - (u_t^\epsilon)_j (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ &\quad + (u_t^\epsilon)_i (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_j (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_i \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho + (u_t^\epsilon)_j \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho + \Sigma_{ij}(t, x_t^\epsilon) dt, \end{aligned} \quad (3.10)$$

where we employed the equation for u_t^ϵ , Eq. (3.1), and Eq. (3.9). We isolate the u -dependent terms that appear in $(u_t^\epsilon)_j \partial_{p_l} H^\epsilon(t, x_t^\epsilon) dt$ to find

$$\begin{aligned} &\frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) (\tilde{\gamma}_{jk}(t, q_t^\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l (u_t^\epsilon)_i \\ &\quad + \tilde{\gamma}_{ik}(t, q_t^\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l (u_t^\epsilon)_j) dt \\ &= -d((u_t^\epsilon)_i (u_t^\epsilon)_j) - (u_t^\epsilon)_i (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt - (u_t^\epsilon)_j (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ &\quad + (u_t^\epsilon)_i (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_j (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_i \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho + (u_t^\epsilon)_j \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho + \Sigma_{ij}(t, x_t^\epsilon) dt. \end{aligned} \quad (3.11)$$

We will solve this equation for $\tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(u_t^\epsilon)_j(u_t^\epsilon)_k dt$ using a Lyapunov equation technique, as in [19, 21].

The formula for the left-hand side of Eq. (3.11) clearly represents a differential of a C^1 -function, so, the integral from 0 to t of the right-hand side, which we denote by $(C_t)_{ij}$ is a C^1 -function P -a.s. Differentiating both sides with respect to t , we obtain

$$\begin{aligned} & \frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(A\tilde{\gamma})_j^l(t, q_t^\epsilon)(u_t^\epsilon)_l(u_t^\epsilon)_i \\ & + \frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(A\tilde{\gamma})_i^l(t, q_t^\epsilon)(u_t^\epsilon)_l(u_t^\epsilon)_j = (\dot{C}_t)_{ij}, \end{aligned} \quad (3.12)$$

where we define $(A\tilde{\gamma})_j^i = \tilde{\gamma}_{jk}A^{ki}$.

Defining the matrix

$$(V_t)_{ij} = \frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(u_t^\epsilon)_i(u_t^\epsilon)_j \quad (3.13)$$

we rewrite Eq. (3.12) as

$$(A\tilde{\gamma})_i^l V_{lj} + V_{il} (A\tilde{\gamma})_j^l = \dot{C}_{ij}. \quad (3.14)$$

This is a Lyapunov equation for V .

For every $T > 0$, there exists $c > 0$ and $\lambda > 0$ such that $-A\tilde{\gamma}$ has eigenvalues with real parts bounded above by $-c\lambda$, uniformly on $[0, T] \times \mathbb{R}^n$. See Lemma A2. Hence, we can solve uniquely for V ,

$$V_{ij} = \int_0^\infty (e^{-yA\tilde{\gamma}})_i^k \dot{C}_{kl} (e^{-yA\tilde{\gamma}})_j^l dy. \quad (3.15)$$

See, for example, Theorem 6.4.2 in [25].

Remark 2. *Assumption 5, or something else that accomplishes a similar purpose, is necessary in the above computation. If one tries to solve for $(\tilde{V}_t)_j^i \equiv (u_t^\epsilon)_j \partial_{p_i} H^\epsilon(t, x_t^\epsilon)$ directly, then one is led to the linear equation*

$$\tilde{\gamma}_{jk} \tilde{V}_i^k + \tilde{\gamma}_{ik} \tilde{V}_j^k = \dot{C}_{ij}. \quad (3.16)$$

The left-hand side of this equation has a non-trivial kernel, consisting of all \tilde{V} for which $\tilde{\gamma}_{ik} \tilde{V}_j^k$ is antisymmetric. Therefore, just knowing that \tilde{V} satisfies Eq. (3.16) does not allow us to uniquely solve for \tilde{V} . Some additional constraint must be combined with Eq. (3.16) in order to solve for \tilde{V} .

Integrating Eq. (3.15) with respect to time, we obtain

$$\begin{aligned} & \frac{2}{\epsilon} \int_0^t \tilde{K}'(s, q_s^\epsilon, \|u_s^\epsilon\|_A^2/\epsilon) (u_s^\epsilon)_i (u_s^\epsilon)_j ds \\ &= \int_0^t \int_0^\infty \left(e^{-y(A\tilde{\gamma})(s, q_s^\epsilon)} \right)_i^k \left(e^{-y(A\tilde{\gamma})(s, q_s^\epsilon)} \right)_j^l dy (\dot{C}_s)_{kl} ds. \end{aligned} \quad (3.17)$$

The functions

$$G_{ij}^{kl}(t, q) = \int_0^\infty \left(e^{-y(A\tilde{\gamma})(t, q)} \right)_i^k \left(e^{-y(A\tilde{\gamma})(t, q)} \right)_j^l dy \quad (3.18)$$

are C^1 , hence $G_{ij}^{kl}(t, q_t^\epsilon)$ are semimartingales and

$$\begin{aligned} & \frac{2}{\epsilon} \tilde{K}'(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) (u_t^\epsilon)_i (u_t^\epsilon)_j dt = G_{ij}^{ab}(t, q_t^\epsilon) d(C_t)_{ab} \\ &= G_{ij}^{ab}(t, q_t^\epsilon) \Sigma_{ab}(t, x_t^\epsilon) dt - G_{ij}^{ab}(t, q_t^\epsilon) d((u_t^\epsilon)_a (u_t^\epsilon)_b) \\ & \quad - G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a (\partial_{q^b} K)^\epsilon(t, x_t^\epsilon) dt - G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b (\partial_{q^a} K)^\epsilon(t, x_t^\epsilon) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a (-\partial_t \psi_b(t, q_t^\epsilon) - \partial_{q^b} V(t, q_t^\epsilon) + F_b(t, x_t^\epsilon)) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b (-\partial_t \psi_a(t, q_t^\epsilon) - \partial_{q^a} V(t, q_t^\epsilon) + F_a(t, x_t^\epsilon)) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a \sigma_{b\rho}(t, x_t^\epsilon) dW_t^\rho + G_{ij}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b \sigma_{a\rho}(t, x_t^\epsilon) dW_t^\rho. \end{aligned} \quad (3.19)$$

Combining Eq. (3.6) with Eq. (3.8), Eq. (3.9), and Eq. (3.19) we see that q_t^ϵ satisfies the equation

$$\begin{aligned} d(q_t^\epsilon)^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ & \quad + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) dt \\ & \quad + Q^{ikl}(t, q_t^\epsilon) J_{kl}(t, x_t^\epsilon) dt + d(R_t^\epsilon)^i, \end{aligned} \quad (3.20)$$

where

$$J_{ij}(t, x) \equiv G_{ij}^{kl}(t, q) \Sigma_{kl}(t, x), \quad (3.21)$$

$$Q^{ijl}(t, q) \equiv \partial_{q^k} (\tilde{\gamma}^{-1})^{ij}(t, q) A^{kl}(t, q) - \frac{1}{2} (\tilde{\gamma}^{-1})^{ik}(t, q) \partial_{q^k} A^{jl}(t, q), \quad (3.22)$$

and

$$\begin{aligned}
d(R_t^\epsilon)^i &\equiv -d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\
&\quad - Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) d((u_t^\epsilon)_a (u_t^\epsilon)_b) \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a (-\partial_t \psi_b(t, q_t^\epsilon) - \partial_{q^b} V(t, x_t^\epsilon) \\
&\quad \quad - (\partial_{q^b} K)^\epsilon(t, x_t^\epsilon) + F_b(t, x_t^\epsilon)) dt \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b (-\partial_t \psi_a(t, q_t^\epsilon) - \partial_{q^a} V(t, x_t^\epsilon) \\
&\quad \quad - (\partial_{q^a} K)^\epsilon(t, x_t^\epsilon) + F_a(t, x_t^\epsilon)) dt \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a \sigma_{b\rho}(t, x_t^\epsilon) dW_t^\rho \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b \sigma_{a\rho}(t, x_t^\epsilon) dW_t^\rho.
\end{aligned} \tag{3.23}$$

Based on our knowledge of the decay rate of u_t^ϵ , we expect R_t^ϵ to go to zero in the limit $\epsilon \rightarrow 0^+$. In general, one would still need to extract the portion of $(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(t, q_t^\epsilon, \|z\|_A^2/\epsilon) dt$ that survives in the limit. However, in this paper we will assume:

Assumption 6. $\tilde{K} = \tilde{K}(t, z)$ i.e. \tilde{K} is independent of q , and hence

$$(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(t, q_t^\epsilon, \|z\|_A^2/\epsilon) dt = 0. \tag{3.24}$$

Along with Lemma 2.5, the above calculations motivate the proposed limiting equation

$$\begin{aligned}
dq_t^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t) (-\partial_t \psi_j(t, q_t) - \partial_{q^j} V(t, q_t) + F_j(t, q_t, \psi(t, q_t))) dt \\
&\quad + Q^{ikl}(t, q_t) J_{kl}(t, q_t, \psi(t, q_t)) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t, \psi(t, q_t)) dW_t^\rho.
\end{aligned} \tag{3.25}$$

Note that an additional *noise induced drift* term,

$$S^i(t, q) \equiv Q^{ijl}(t, q) J_{jl}(t, q, \psi(t, q)), \tag{3.26}$$

arises in the limit when Σ is nonzero and (generally) when $\tilde{\gamma}$ and/or A have nontrivial q -dependence. This is in addition to the forcing term, $-\partial_t \psi - \nabla_q V + F$, and is another manifestation of the phenomenon derived in [19, 21].

Remark 3. *Assumption 6 determines the splitting of the Hamiltonian into $K(t, q, p)$ and $V(t, q)$, up to a function of time i.e. if $H = K_1 + V_1 = K_2 + V_2$ are two splittings then $V_1(t, q) = V_2(t, q) + c(t)$. This ambiguity does not impact the limiting equation Eq. (3.25), and so the limiting equation is uniquely defined by the original SDE, Eq. (1.11)-Eq. (1.12), as it has to be, of course.*

4. Convergence Proof

In this final section, we prove convergence of q_t^ϵ to the solution of the proposed limiting equation, Eq. (3.25). This will be accomplished by using the following lemma.

Lemma 4.1. *Let $T > 0$ and suppose we have continuous functions $\tilde{F}(t, x) : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\tilde{\sigma}(t, x) : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times k}$, and $\psi : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ that are Lipschitz in x , uniformly in $t \in [0, T]$.*

Let W_t be a k -dimensional Wiener process, $p \geq 2$ and $\beta > 0$ and suppose that we have continuous semimartingales q_t and, for each $0 < \epsilon \leq \epsilon_0$, R_t^ϵ , $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$ that satisfy the following properties:

1. $q_t^\epsilon = q_0^\epsilon + \int_0^t \tilde{F}(s, x_s^\epsilon) ds + \int_0^t \tilde{\sigma}(s, x_s^\epsilon) dW_s + \tilde{R}_t^\epsilon.$
2. $q_t = q_0 + \int_0^t \tilde{F}(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{\sigma}(s, q_s, \psi(s, q_s)) dW_s.$
3. $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^\beta)$ as $\epsilon \rightarrow 0^+.$
4. $E[\sup_{t \in [0, T]} \|\tilde{R}_t^\epsilon\|^p] = O(\epsilon^\beta)$ as $\epsilon \rightarrow 0^+.$
5. $\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^\beta)$ as $\epsilon \rightarrow 0^+.$
6. $E[\sup_{t \in [0, T]} \|q_t^\epsilon\|^p] < \infty$ for all $\epsilon > 0$ sufficiently small.
7. $E[\sup_{t \in [0, T]} \|q_t\|^p] < \infty.$

Then

$$E \left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (4.1)$$

Proof. The equations for q_t^ϵ and q_t imply

$$\begin{aligned} (q_t^\epsilon)^i - (q_t)^i &= (q_0^\epsilon)^i - (q_0)^i + \int_0^t \tilde{F}^i(s, x_s^\epsilon) - \tilde{F}^i(s, q_s, \psi(s, q_s)) ds \\ &+ \int_0^t [\tilde{\sigma}_\rho^i(s, x_s^\epsilon) - \tilde{\sigma}_\rho^i(s, q_s, \psi(s, q_s))] dW_s^\rho + (\tilde{R}_t^\epsilon)^i. \end{aligned} \quad (4.2)$$

Using properties 3 and 4, along with the Hölder and Burkholder-Davis-Gundy inequalities, for $0 \leq t \leq T$ we have

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \\ & \leq 4^{p-1} \left(E[\|q_0^\epsilon - q_0\|^p] + E \left[\left(\int_0^t \|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\| ds \right)^p \right] \right) \end{aligned} \quad (4.3)$$

$$\begin{aligned} & + E \left[\sup_{s \in [0, t]} \left\| \int_0^s \tilde{\sigma}_\rho^i(r, x_r^\epsilon) - \tilde{\sigma}_\rho^i(r, q_r, \psi(r, q_r)) dW_r^\rho \right\|^p \right] + E \left[\sup_{s \in [0, t]} \|\tilde{R}_s^\epsilon\|^p \right] \\ & \leq 4^{p-1} \left(T^{p-1} E \left[\int_0^t \|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\|^p ds \right] \right) \end{aligned} \quad (4.4)$$

$$\begin{aligned} & + \tilde{C} E \left[\left(\int_0^t \|\tilde{\sigma}(r, x_r^\epsilon) - \tilde{\sigma}(r, q_r, \psi(r, q_r))\|_F^2 dr \right)^{p/2} \right] + O(\epsilon^\beta) \\ & \leq 4^{p-1} \left(T^{p-1} \int_0^t E[\|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\|^p] ds \right) \end{aligned} \quad (4.5)$$

$$+ \tilde{C} T^{p/2-1} \int_0^t E[\|\tilde{\sigma}(r, x_r^\epsilon) - \tilde{\sigma}(r, q_r, \psi(s, q_s))\|_F^p] dr + O(\epsilon^\beta).$$

By assumption, $\tilde{\sigma}$, \tilde{F} , and ψ are Lipschitz in x , uniformly on $[0, T]$. Hence, using property 5,

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p + \|p_s^\epsilon - \psi(s, q_s)\|^p] ds + O(\epsilon^\beta) \quad (4.6) \\ & \leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p + \|p_s^\epsilon - \psi(s, q_s^\epsilon)\|^p + \|\psi(s, q_s^\epsilon) - \psi(s, q_s)\|^p] ds + O(\epsilon^\beta) \\ & \leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p] ds + \tilde{C} \sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] + O(\epsilon^\beta) \\ & = \tilde{C} \int_0^t E \left[\sup_{r \in [0, s]} \|q_r^\epsilon - q_r\|^p \right] ds + O(\epsilon^\beta) \end{aligned}$$

for all $0 \leq t \leq T$, where the constants change from line to line and are all independent of t .

Properties 6 and 7 imply that $E[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p] \in L^1([0, T])$ for ϵ sufficiently small, and hence Gronwall's inequality applied to Eq. (4.6) gives

$$E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \leq O(\epsilon^\beta) e^{\tilde{C}t} \quad (4.7)$$

for $0 \leq t \leq T$. □

To prove that the hypotheses of Lemma 4.1 hold, we will need the following assumption:

Assumption 7. *We assume that, for every $T > 0$, $\nabla_q V$, F , and σ are Lipschitz in x uniformly in $t \in [0, T]$. We also assume that A and γ are C^2 , ψ is C^3 , and $\partial_t \psi$, $\partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \psi$, $\partial_t \partial_{q^i} \psi$, $\partial_t \partial_{q^j} \partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \partial_{q^i} \psi$, $\partial_t \gamma$, $\partial_{q^i} \gamma$, $\partial_t \partial_{q^j} \gamma$, $\partial_{q^i} \partial_{q^j} \gamma$, $\partial_t A$, $\partial_{q^i} A$, $\partial_t \partial_{q^i} A$, and $\partial_{q^i} \partial_{q^j} A$ are bounded on $[0, T] \times \mathbb{R}^{2n}$ for every $T > 0$.*

Note that, combined with our prior assumptions, this implies $\tilde{\gamma}$, $\tilde{\gamma}^{-1}$, $\partial_t \tilde{\gamma}^{-1}$, $\partial_{q^i} \tilde{\gamma}^{-1}$, $\partial_t \partial_{q^j} \tilde{\gamma}^{-1}$, and $\partial_{q^i} \partial_{q^j} \tilde{\gamma}^{-1}$ are bounded on compact t intervals. Additionally, using the formula for the derivative of the matrix exponential found in [26], one can prove that our assumptions also imply that the G_{kl}^{ij} 's are bounded and Lipschitz in q , uniformly on compact t intervals.

As a step towards using Lemma 4.1 to prove our convergence result, we now show that R_t^ϵ from Eq. (3.23) converges to zero in the appropriate sense.

Lemma 4.2. *Under Assumptions 1-5, 7, for any $p > 0$, $T > 0$, $0 < \beta < p/2$ we have*

$$E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+, \quad (4.8)$$

where R_t^ϵ was defined in Eq. (3.23).

Proof. Let us first assume that $p > 2$. Let $0 < \beta < p/2$. Define

$$Y(t, x) = -\partial_t \psi(t, q) - \nabla_q V(t, q) + F(t, x). \quad (4.9)$$

Our assumptions imply that Y is bounded on $[0, T] \times \mathbb{R}^{2n}$.

From Eq. (3.23),

$$\begin{aligned}
E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] &\leq 8^{p-1} \left(E \left[\sup_{t \in [0, T]} \|(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j\|^p \right] \right. \\
&+ E[\|(\tilde{\gamma}^{-1})^{ij}(0, q_0^\epsilon)(u_0^\epsilon)_j\|^p] + E \left[\left(\int_0^T \|(u_s^\epsilon)_j \partial_s (\tilde{\gamma}^{-1})^{ij}(s, q_s^\epsilon)\| ds \right)^p \right] \\
&+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) d((u_s^\epsilon)_a (u_s^\epsilon)_b) \right\|^p \right] \\
&+ E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\
&+ E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_b (-\partial_{q^a} K)^\epsilon(s, x_s^\epsilon) + Y_a(s, x_s^\epsilon)\| ds \right)^p \right] \\
&+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \\
&+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_b \sigma_{a\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \Big)
\end{aligned} \tag{4.10}$$

where the norm is the 2-norm of vectors, with components indexed by i , resulting from summation over other, repeated indices. We now show that all of these terms are $O(\epsilon^\beta)$.

Boundedness of $\tilde{\gamma}^{-1}$ together with Lemma 2.5 implies that the first two terms satisfy

$$\begin{aligned}
&E \left[\sup_{t \in [0, T]} \|(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j\|^p \right] + E[\|(\tilde{\gamma}^{-1})^{ij}(0, q_0^\epsilon)(u_0^\epsilon)_j\|^p] \\
&\leq 2\|\tilde{\gamma}^{-1}\|_\infty^p E \left[\sup_{t \in [0, T]} \|u_t^\epsilon\|^p \right] = O(\epsilon^\beta).
\end{aligned} \tag{4.11}$$

By Hölder's inequality, boundedness of $\partial_t \tilde{\gamma}^{-1}$, and Lemma 2.5, the third term satisfies

$$E \left[\left(\int_0^T \|(u_s^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(s, q_s^\epsilon)\| ds \right)^p \right] \tag{4.12}$$

$$\leq T^{p-1} \|\partial_t \tilde{\gamma}^{-1}\|_\infty^p \int_0^T E[\|u_s^\epsilon\|^p] ds \tag{4.13}$$

$$\leq T^p \|\partial_t \tilde{\gamma}^{-1}\|_\infty^p \sup_{s \in [0, T]} E[\|u_s^\epsilon\|^p] = O(\epsilon^{p/2}).$$

The functions $Q^{ijl}(t, q)G_{jl}^{ab}(t, q)$ are C^1 , bounded, with bounded first derivatives on $[0, T] \times \mathbb{R}^n$. Therefore, by Proposition 2.3 we have

$$E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) d((u_s^\epsilon)_a (u_s^\epsilon)_b) \right\|^p \right] = O(\epsilon^{p/2}). \quad (4.14)$$

Using Hölder's inequality, Assumption 1, and our various boundedness assumptions, the fifth term can be bounded as follows:

$$\begin{aligned} & E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\ & \leq \tilde{C} T^{p-1} E \left[\int_0^T \|u_s^\epsilon\|^p (\|(\partial_{q^b} K)^\epsilon(s, x_s^\epsilon)\| + \|Y\|_\infty)^p ds \right] \end{aligned} \quad (4.15)$$

$$\leq \tilde{C} T^p \sup_{s \in [0, T]} E [\|u_s^\epsilon\|^p (M + C K^\epsilon(s, x_s^\epsilon) + \|Y\|_\infty)^p]. \quad (4.16)$$

Again, here and in the following, we let \tilde{C} denote a constant that may vary from line to line. We now use Cauchy-Schwarz inequality, Lemma 2.5 and Proposition 2.1 to obtain

$$\begin{aligned} & E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\ & \leq \tilde{C} T^p \sup_{s \in [0, T]} E [\|u_s^\epsilon\|^{2p}]^{1/2} \sup_{s \in [0, T]} E [(M + C K^\epsilon(s, x_s^\epsilon) + \|Y\|_\infty)^{2p}]^{1/2} \\ & = O(\epsilon^p)^{1/2} O(1) = O(\epsilon^{p/2}). \end{aligned} \quad (4.17)$$

A similar argument shows that the sixth term is also $O(\epsilon^{p/2})$.

Using the Burkholder-Davis-Gundy and Hölder inequalities together with the boundedness assumptions and Lemma 2.5, the seventh term satisfies

$$\begin{aligned} & E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \\ & \leq \tilde{C} E \left[\left(\int_0^T \sum_{i, \rho} (Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon))^2 ds \right)^{p/2} \right] \\ & \leq \tilde{C} E \left[\left(\int_0^T \|u_s^\epsilon\|^2 ds \right)^{p/2} \right] \leq \tilde{C} T^{p/2-1} E \left[\int_0^T \|u_s^\epsilon\|^p ds \right] \\ & \leq \tilde{C} \sup_{s \in [0, T]} E [\|u_s^\epsilon\|^p] = O(\epsilon^{p/2}), \end{aligned} \quad (4.18)$$

where the power of T can be absorbed into the constant, since we are working on a fixed time interval. A similar estimate applies to the final, eighth term. Therefore, we have proven the claim for $p > 2$. An application of Hölder's inequality proves it for all $p > 0$. \square

We now have all the ingredients to prove convergence of q_t^ϵ to q_t .

Theorem 4.1. *Let x_t^ϵ be a family of solutions to the SDE Eq. (1.11)-Eq. (1.12) with initial condition $(q_0^\epsilon, p_0^\epsilon)$ and q_t be a solution to the proposed limiting SDE, Eq. (3.25), with initial condition q_0 . Suppose that for all $\epsilon > 0$ and all $p > 0$ we have $E[\|q_0^\epsilon\|^p] < \infty$, $E[\|q_0\|^p] < \infty$, and $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$. Also suppose that Assumptions 1-7 hold. Then for any $T > 0$, $p > 0$, $0 < \beta < p/2$ we have*

$$E \left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (4.19)$$

Proof. First let $p > 2$. Define the generalized force vector field

$$\tilde{F}(t, x)^i = (\tilde{\gamma}^{-1})^{ij}(t, q)(-\partial_t \psi_i(t, q) - \partial_{q^j} V(t, q) + F_j(t, x)) + Q^{ijl}(t, q) J_{jl}(t, x) \quad (4.20)$$

and noise coefficients

$$\tilde{\sigma}_\eta^i(t, x) = (\tilde{\gamma}^{-1})^{ij}(t, q) \sigma_{j\eta}(t, x). \quad (4.21)$$

Our assumptions imply that these are bounded on $[0, T] \times \mathbb{R}^{2n}$ and, along with ψ , they are Lipschitz in x , uniformly in $t \in [0, T]$.

We will now check all of the properties that are required to use Lemma 4.1. By Eq. (3.20) and Eq. (3.25), q_t^ϵ and q_t satisfy the equations

$$(q_t^\epsilon)^i = (q_0^\epsilon)^i + \int_0^t \tilde{F}^i(s, x_s^\epsilon) ds + \int_0^t \tilde{\sigma}_\eta^i(s, x_s^\epsilon) dW_s^\eta + (R_t^\epsilon)^i \quad (4.22)$$

and

$$(q_t)^i = (q_0)^i + \int_0^t \tilde{F}^i(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{\sigma}_\eta^i(s, q_s, \psi(s, q_s)) dW_s^\eta. \quad (4.23)$$

Note, that the term involving $\partial_{q^j} \tilde{K}$, present in Eq. (3.20), vanishes under Assumption 6. In addition, for each $T > 0$, $\tilde{F}(t, q, \psi(t, q))$ and $\tilde{\sigma}(t, q, \psi(t, q))$ are bounded and Lipschitz in q , uniformly in t on $[0, T] \times \mathbb{R}^n$, so a unique solution to Eq. (4.23) exists and is defined for all $t \geq 0$ [27].

By assumption, the initial conditions satisfy $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2}) = O(\epsilon^\beta)$. Lemma 4.2 implies that

$$E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+ \quad (4.24)$$

and, by Lemma 2.5,

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (4.25)$$

For any $\epsilon > 0$, $p > 1$ we have

$$\begin{aligned} E \left[\sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] &= E \left[\sup_{t \in [0, T]} \left\| q_0^\epsilon + \int_0^t \nabla_p H^\epsilon(r, x_r^\epsilon) dr \right\|^p \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} E \left[\left(\int_0^T \|\nabla_p K^\epsilon(r, x_r^\epsilon)\| dr \right)^p \right] \\ &= 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} E \left[\left(\int_0^T \epsilon^{-1/2} \|(\nabla_z K)^\epsilon(r, x_r^\epsilon)\| dr \right)^p \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} T^{p-1} \epsilon^{-p/2} E \left[\int_0^T (M + K^\epsilon(r, x_r^\epsilon))^p dr \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 4^{p-1} T^p \epsilon^{-p/2} (M^p + \sup_{r \in [0, T]} E[K^\epsilon(r, x_r^\epsilon)^p]) < \infty, \end{aligned} \quad (4.26)$$

where we used Assumption 1 to bound $\nabla_z K$ and Proposition 2.1 in the last line.

We also have

$$\begin{aligned} E \left[\sup_{t \in [0, T]} \|q_t\|^p \right] &\leq 3^{p-1} \left(E[\|q_0\|^p] + E \left[\left(\int_0^T \|\tilde{F}(s, q_s, \psi(s, q_s))\| ds \right)^p \right] \right. \\ &\quad \left. + E \left[\sup_{t \in [0, T]} \left\| \int_0^t \tilde{\sigma}(s, q_s, \psi(s, q_s)) dW_s \right\|^p \right] \right). \end{aligned} \quad (4.27)$$

\tilde{F} and $\tilde{\sigma}$ are bounded uniformly up to time T , so applying the Burkholder-Davis-Gundy and Hölder inequalities to the last term, we get

$$E \left[\sup_{t \in [0, T]} \|q_t\|^p \right] \leq 3^{p-1} (E[\|q_0\|^p] + T^p \|\tilde{F}\|_\infty^p + \tilde{C} T^{p/2} \|\tilde{\sigma}\|_{E, \infty}^p) < \infty. \quad (4.28)$$

Hence we have verified properties 1-7 from Lemma 4.1 for every $T > 0$, $p > 2$, and $0 < \beta < p/2$, and can conclude the convergence result, Eq. (4.19). The result for any $p > 0$ follows from an application of Hölder's inequality. \square

Corollary 4.1. *In particular, given the above assumptions on the initial conditions, γ , σ , F , etc., Theorem 4.1 holds for Hamiltonians of the form*

$$H(t, q, p) = \sum_{l=k_1}^{k_2} d_l(t) \left[A^{ij}(t, q) (p - \psi(t, q))_i (p - \psi(t, q))_j \right]^l + V(t, q) \quad (4.29)$$

where $1 \leq k_1 \leq k_2$ are integers and the following properties hold on $[0, T] \times \mathbb{R}^n$ for every $T > 0$:

1. V is C^2 and $\nabla_q V$ is bounded and Lipschitz in q , uniformly in t .
2. ψ is C^3 and $\partial_t \psi$, $\partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \psi$, $\partial_t \partial_{q^i} \psi$, $\partial_t \partial_{q^j} \partial_{q^i} \psi$, and $\partial_{q^i} \partial_{q^j} \partial_{q^i} \psi$ are bounded.
3. d_l are C^2 and non-negative.
4. d_{k_1} and d_{k_2} are uniformly bounded below by a positive constant.
5. A is C^2 , positive-definite, and A , $\partial_t A$, $\partial_{q^i} A$, $\partial_t \partial_{q^i} A$, and $\partial_{q^i} \partial_{q^j} A$ are bounded.
6. The eigenvalues of A are uniformly bounded below by a positive constant.

Verification of the assumptions of Theorem 4.1 is elementary and we leave it to the reader.

5. Examples

In this section, we elaborate on two important examples that are covered by the framework we have developed.

5.1. Particle in a Electromagnetic Field

The Hamiltonian of a particle of mass m and charge e in an electromagnetic field with vector potential $\phi(t, q)$ and electrostatic potential $V(t, q)$ is

$$H(t, q, p) = \frac{1}{2m} \|p - e\phi(t, q)\|^2 + eV(t, q). \quad (5.1)$$

Allowing for an additional forcing, F , Hamilton's equations for this system are

$$dq_t^\epsilon = \frac{1}{\epsilon m} (p_t^\epsilon - e\phi(t, q_t^\epsilon)) dt, \quad (5.2)$$

$$d(p_t^\epsilon)_i = \left(-\frac{1}{\epsilon m} \gamma_{ij}(t, q_t^\epsilon) \delta^{jk} ((p_t^\epsilon)_k - e\phi_k(t, q_t^\epsilon)) + F_i(t, x_t^\epsilon) - e\partial_{q^i} V(t, q) \right. \\ \left. - \frac{e}{\epsilon m} \partial_{q^i} \phi_k(t, q_t^\epsilon) \delta^{jk} ((p_t^\epsilon)_j - e\phi_j(t, q_t^\epsilon)) \right) dt + \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho. \quad (5.3)$$

The homogenized equation in the small mass limit, Eq. (3.25), is difficult to simplify in general. However, in the case where γ and σ are independent of p and the fluctuation dissipation relation holds pointwise for a time and position dependent “temperature” $T(t, q)$,

$$\Sigma_{ij}(t, q) = 2k_B T(t, q) \gamma_{ij}(t, q), \quad (5.4)$$

one can show that

$$G_{kl}^{ab}(t, q) \Sigma_{ab}(t, q) = k_B T(t, q) \delta_{kl}, \quad (5.5)$$

where G was defined in Eq. (3.18) and $\Sigma_{ij} = \sum_\rho \sigma_{i\rho} \sigma_{j\rho}$.

The noise induced drift, Eq. (3.26), can therefore be simplified to

$$S^i(t, q) = k_B T(t, q) \partial_{q^j} (\tilde{\gamma}^{-1})(t, q_t)^{ij}. \quad (5.6)$$

Recall that we defined

$$\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q), \quad (5.7)$$

where here, $\psi = e\phi$.

The homogenized equation in the small mass limit is then

$$dq_t^i = (\tilde{\gamma}^{-1})^{ij}(t, q_t) (-\partial_t \psi_j(t, q_t) - e\partial_{q^j} V(t, q_t) + F_j(t, q_t, \psi(t, q_t))) dt \\ + k_B T(t, q) \partial_{q^j} (\tilde{\gamma}^{-1})(t, q_t)^{ij} dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t) dW_t^\rho. \quad (5.8)$$

This agrees with the result obtained in [19].

5.2. Particle on a Riemannian Manifold

Another case that is covered by the framework developed in this paper is the inertial motion of a particle in \mathbb{R}^n , but with geometry specified by a Riemannian metric tensor, $g_{ij}(t, q)$. The Hamiltonian of such a system is

$$H(t, q, p) = \frac{1}{2m} g^{ij}(t, q) p_i p_j. \quad (5.9)$$

Note that the inverse metric tensor, $g^{ij}(t, q)$, is playing the role of $A^{ij}(t, q)$ in the formalism developed in the previous sections, and so all the assumptions that were required of A^{ij} there must be satisfied by g^{ij} here.

Allowing for external forcing, F , Hamilton's equations are

$$d(q_t^\epsilon)^i = \frac{1}{\epsilon m} g^{ij}(t, q_t^\epsilon) (p_t^\epsilon)_j dt, \quad (5.10)$$

$$d(p_t^\epsilon)_i = \left(-\frac{1}{\epsilon m} \gamma_{ij}(t, x_t^\epsilon) g^{jk}(t, q_t^\epsilon) (p_t^\epsilon)_k - \frac{1}{2\epsilon m} \partial_{q_i} g^{kl}(t, q_t^\epsilon) (p_t^\epsilon)_k (p_t^\epsilon)_l + F_i(t, x_t^\epsilon) \right) dt \quad (5.11)$$

$$+ \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho.$$

Again, the homogenized equation, Eq. (3.25), can be simplified if γ and σ are independent of p and the fluctuation dissipation relation holds pointwise for a time and position dependent “temperature” $T(t, q)$,

$$\Sigma_{ij}(t, q) = 2k_B T(t, q) \gamma_{ij}(t, q). \quad (5.12)$$

In this case one finds that

$$G_{kl}^{ab}(t, q) \Sigma_{ab}(t, q) = k_B T(t, q) g_{kl}(t, q) \quad (5.13)$$

and hence Eq. (3.25) becomes

$$dq_t^i = (\gamma^{-1})^{ij}(t, q_t) F_j(t, q_t, 0) dt + S^i(t, q_t) dt + (\gamma^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t) dW_t^\rho, \quad (5.14)$$

where the noise induced drift is

$$S^i(t, q) = k_B T(t, q) \left(\partial_{q^j} (\gamma^{-1})^{ij}(t, q) - \frac{1}{2} (\gamma^{-1})^{ij}(t, q) g_{kl}(t, q) \partial_{q^j} g^{kl}(t, q) \right). \quad (5.15)$$

See also [21], which treats the case of a smooth, compact, connected, manifold without boundary (but otherwise arbitrary topology) via a more geometrically motivated approach.

Appendix A. Linear Algebra Lemmas

Lemma A1. *Let A be an $n \times n$ real or complex matrix with symmetric part $A^s = \frac{1}{2}(A + A^*)$. If the eigenvalues of A^s are bounded above (resp. below) by α then the real parts of the eigenvalues of A are bounded above (resp. below) by α .*

Proof. Suppose y is an eigenvector of A with norm 1 corresponding to the eigenvalue λ . Let A^a be the antisymmetric part of A .

$$\Re(y^* A^a y) = \Re(\overline{y^* A^a y}) = \Re(y^* (A^a)^* y) = -\Re(y^* A^a y). \quad (\text{A.1})$$

Therefore $\Re(y^* A^a y) = 0$ and

$$\Re(\lambda) = \Re(y^* A y) = \Re(y^* A^s y). \quad (\text{A.2})$$

If the eigenvalues of A^s are bounded above by α then

$$\Re(\lambda) = \Re(y^* A^s y) \leq \alpha \|y\|^2 = \alpha \quad (\text{A.3})$$

and if they are bounded below by α then

$$\Re(\lambda) = \Re(y^* A^s y) \geq \alpha \|y\|^2 = \alpha. \quad (\text{A.4})$$

□

Lemma A2. *Let A be a positive definite $n \times n$ -real matrix with eigenvalues bounded below by $\lambda > 0$ and B be an $n \times n$ -real matrix whose symmetric part has eigenvalues bounded below by $\gamma > 0$. Then the eigenvalues of AB have real parts bounded below by $\gamma\lambda$.*

Proof. A is positive definite, so we can factor it as $A = DD^T$ where D is a real valued, invertible, $n \times n$ -matrix. AB and the conjugation $D^{-1}ABD = D^TBD$ have the same eigenvalues. The symmetric part of D^TBD is D^TB^sD and for any $y \in \mathbb{R}^n$,

$$y^T D^T B^s D y \geq \gamma y^T D^T D y. \quad (\text{A.5})$$

$D^T D$ is a positive definite matrix that has the same eigenvalues as $A = DD^T$ (both are equal to the squared singular values of D). The eigenvalues of A are bounded below by λ , so

$$y^T D^T D y \geq \lambda \|y\|^2. \quad (\text{A.6})$$

Therefore the eigenvalues of the symmetric part of D^TBD are bounded below by $\gamma\lambda$. Hence, by Lemma A1, the real parts of the eigenvalues of D^TBD , and hence of AB , are bounded below by $\gamma\lambda$. □

Appendix B. Non-Explosion of Solutions

In the course of proving our main result, Theorem 4.1, we showed that the limiting process, q_t , exist for all $t \geq 0$ with probability one. Though we have assumed it to be the case throughout this paper, the same is not obvious for the solutions, x_t^ϵ , to the SDE Eq. (1.11)-Eq. (1.12). However, existence for all $t \geq 0$ can be proven under a collection of assumptions that are very similar to our Assumptions 1-7 from the main text, as shown in the following lemma. We emphasize that in this appendix, we do not employ any of the Assumptions 1-7 per se. The assumptions we do use are all listed below, in the statement of the lemma.

Lemma B1. *Suppose:*

1. $H(t, x)$ has the form given in Eq. (1.8) and the subsequent text.
2. There exist $C > 0$ and $M > 0$ such that

$$\max \left\{ |\partial_t K(t, q, z)|, \left(\sum_{ij} |\partial_{z_i} \partial_{z_j} K(t, q, z)|^2 \right)^{1/2} \right\} \leq M + CK(t, q, z). \quad (\text{B.1})$$

3. For every $t \geq 0$ there exists $c > 0$, $\eta \geq 1$ such that

$$K(t, q, z) \geq c \|z\|^{2\eta}. \quad (\text{B.2})$$

4. $F(t, x)$ is continuous and is locally Lipschitz in x with the Lipschitz constant uniform on compact time intervals.
5. $\sigma(t, x)$ is continuous and bounded and is locally Lipschitz in x with the Lipschitz constant uniform on compact time intervals.
6. $\gamma(t, x)$ is continuous and bounded and is locally Lipschitz in x with the Lipschitz constant uniform on compact time intervals. Also, its eigenvalues (which are real, since γ is symmetric) are bounded below by some $\lambda > 0$.

7. There exists $M > 0$ such that

$$\|\psi(t, q)\|^2 \leq M(1 + \|q\|^2) \quad (\text{B.3})$$

for all t, x . i.e. $\psi(t, x)$ is linearly bounded in q , uniformly in t .

8. $q - \partial_t \psi(t, q) - \nabla_q V(t, q) + F(t, x)$ is linearly bounded in x , uniformly in t .

Then the maximal solutions, x_t^ϵ , to the SDE Eq. (1.11)-Eq. (1.12) are a.s. unique and a.s. exist for all $t \geq 0$.

Proof. Fix $\epsilon > 0$. The assumptions imply that an a.s. unique, maximal solution x_t^ϵ exists up to explosion time e^ϵ (see Section 3.4 in [27]). Non-explosion of x_t^ϵ (i.e. $e^\epsilon = \infty$ a.s.) will follow from the existence of a Lyapunov function (see Theorem 3.5 in [27]), a non-negative C^2 function, $U(t, x)$, that satisfies:

1. For any $t \geq 0$, $\lim_{x \rightarrow \infty} U(t, x) = \infty$,
- 2.

$$L[U](t, x) \leq \tilde{M} + \tilde{C}U(t, x) \quad (\text{B.4})$$

for some $\tilde{M} \geq 0$, $\tilde{C} > 0$, where L is the time dependent generator

$$\begin{aligned} L[U](t, x) = & \partial_t U(t, x) + \nabla_p H^\epsilon(t, x) \cdot \nabla_q U(t, x) \\ & + (-\gamma(t, x) \nabla_p H^\epsilon(t, x) - \nabla_q H^\epsilon(t, x) + F(t, x)) \cdot \nabla_p U(t, x) \\ & + \frac{1}{2} \Sigma_{ij}(t, x) \partial_{p_i} \partial_{p_j} U(t, x) \end{aligned} \quad (\text{B.5})$$

and $\Sigma_{ij} = \sum_\rho \sigma_{i\rho} \sigma_{j\rho}$.

We note that to connect with the result as stated in [27], one needs to work with $\tilde{M}/\tilde{C} + U$ in place of U , but we find the above formulation more convenient here.

As our candidate Lyapunov function, we define

$$U(t, x) = \frac{1}{2} \|q\|^2 + K^\epsilon(t, x). \quad (\text{B.6})$$

K^ϵ is non-negative and C^2 , therefore U is as well.

Fix $t \geq 0$. Given $R > 0$, let

$$\max\{\|q\|, \|p\|\} \geq (2R)^{1/2} + \sup_{\|q\| \leq R} \|\psi(t, q)\| + \epsilon^{1/2} (R/c)^{1/2\eta}. \quad (\text{B.7})$$

Therefore we either have $U(t, x) \geq \frac{1}{2}\|q\|^2 \geq R$ or $\|q\| < (2R)^{1/2}$ and hence

$$\begin{aligned}
U(t, x) &\geq K^\epsilon(t, q, p) \geq c\|p - \psi(t, q)\|^{2\eta}/\epsilon^\eta \\
&\geq \frac{c}{\epsilon^\eta} \left(\|p\| - \sup_{\|q\| \leq R} \|\psi(t, q)\| \right)^{2\eta} \\
&\geq \frac{c}{\epsilon^\eta} \left(\epsilon^{\frac{1}{2}} \left(\frac{R}{c} \right)^{\frac{1}{2\eta}} \right)^{2\eta} = R.
\end{aligned} \tag{B.8}$$

Hence $U(t, x) \rightarrow \infty$ as $x \rightarrow \infty$.

We now compute

$$\begin{aligned}
&L[U](t, x) \\
&= \partial_t K^\epsilon(t, x) + \nabla_p K^\epsilon(t, x) \cdot \nabla_q K^\epsilon(t, x) + \nabla_p K^\epsilon(t, x) \cdot q
\end{aligned} \tag{B.9}$$

$$\begin{aligned}
&+ (-\gamma(t, x) \nabla_p K^\epsilon(t, x) - \nabla_q K^\epsilon(t, x) - \nabla_q V(t, x) + F(t, x)) \cdot \nabla_p K^\epsilon(t, x) \\
&+ \frac{1}{2} \Sigma_{ij}(t, x) \partial_{p_i} \partial_{p_j} K^\epsilon(t, x) \\
&\leq (\partial_t K)^\epsilon(t, x) - \nabla_p K^\epsilon(t, x) \cdot \partial_t \psi(t, q) + \nabla_p K^\epsilon(t, x) \cdot q
\end{aligned} \tag{B.10}$$

$$\begin{aligned}
&- \lambda \|\nabla_p K^\epsilon(t, x)\|^2 + (-\nabla_q V(t, x) + F(t, x)) \cdot \nabla_p K^\epsilon(t, x) \\
&+ \frac{1}{2\epsilon} \Sigma_{ij}(t, x) (\partial_{z_i} \partial_{z_j} K)^\epsilon(t, x) \\
&\leq M + CK^\epsilon(t, x) - \lambda \|\nabla_p K^\epsilon(t, x)\|^2 \\
&+ \|q - \partial_t \psi(t, q) - \nabla_q V(t, x) + F(t, x)\| \|\nabla_p K^\epsilon(t, x)\| \\
&+ \frac{1}{2\epsilon} \|\Sigma\|_{F, \infty} (M + CK^\epsilon(t, x)).
\end{aligned} \tag{B.11}$$

Note that in the above calculation, the $\nabla_p K^\epsilon(t, x) \cdot \nabla_q K^\epsilon(t, x)$ terms cancel.

Using the inequality

$$ab \leq \frac{1}{2}(\delta a^2 + \frac{1}{\delta} b^2) \tag{B.12}$$

for any a, b and any $\delta > 0$ we have

$$\begin{aligned}
& L[U](t, x) \\
& \leq \left(1 + \frac{1}{2\epsilon} \|\Sigma\|_{F, \infty}\right) (M + CK^\epsilon(t, x)) - \lambda \|\nabla_p K^\epsilon(t, x)\|^2 \\
& \quad + \frac{1}{4\lambda} \|q - \partial_t \psi(t, q) - \nabla_q V(t, x) + F(t, x)\|^2 + \lambda \|\nabla_p K^\epsilon(t, x)\|^2 \\
& \leq \left(1 + \frac{1}{2\epsilon} \|\Sigma\|_{F, \infty}\right) (M + CK^\epsilon(t, x)) + \frac{M}{4\lambda} (1 + \|x\|^2) \\
& \leq \left(1 + \frac{1}{2\epsilon} \|\Sigma\|_{F, \infty}\right) (M + CK^\epsilon(t, x)) + \frac{M}{4\lambda} (1 + \|q\|^2) \\
& \quad + \frac{M}{4\lambda} \|p\|^2.
\end{aligned} \tag{B.13}$$

The first two terms are bounded by $\tilde{M} + \tilde{C}U(t, x)$ for some $\tilde{M} \geq 0$, $\tilde{C} \geq 0$. As for the last term, utilizing point 3 from the statement of the lemma, we find

$$\begin{aligned}
\|p\|^2 & \leq (\|p - \psi(t, q)\| + \|\psi(t, q)\|)^2 \\
& \leq 2(1 + \epsilon^\eta \|(p - \psi(t, q))\|/\sqrt{\epsilon}\|^{2\eta} + \|\psi(t, q)\|^2) \\
& \leq 2(1 + \epsilon^\eta K^\epsilon(t, x)/c + M(1 + \|q\|^2)) \\
& \leq \tilde{M} + \tilde{C}U(t, x).
\end{aligned} \tag{B.14}$$

Therefore U is a Lyapunov function. This proves that $e^\epsilon = \infty$ a.s. i.e. the solution x_t^ϵ exists for all $t \geq 0$ a.s. □

Acknowledgments

J.W. was partially supported by NSF grants DMS 131271 and DMS 1615045.

References

- [1] M. Smoluchowski, “Drei vortrage uber diffusion, brownsche bewegung und koagulation von kolloidteilchen,” *Zeitschrift fur Physik*, vol. 17, pp. 557–585, 1916.
- [2] H. Kramers, “Brownian motion in a field of force and the diffusion model of chemical reactions,” *Physica*, vol. 7, no. 4, pp. 284 – 304, 1940.

- [3] G. A. Pavliotis and A. M. Stuart, “White noise limits for inertial particles in a random field,” *Multiscale Modeling & Simulation*, vol. 1, no. 4, pp. 527–553, 2003.
- [4] C. Chevalier and F. Debbasch, “Relativistic diffusions: A unifying approach,” *Journal of Mathematical Physics*, vol. 49, no. 4, 2008.
- [5] I. Bailleul, “A stochastic approach to relativistic diffusions,” in *Annales de l’institut Henri Poincaré (B)*, vol. 46, pp. 760–795, 2010.
- [6] M. A. Pinsky, “Isotropic transport process on a riemannian manifold,” *Transactions of the American Mathematical Society*, vol. 218, pp. 353–360, 1976.
- [7] M. A. Pinsky, “Homogenization in stochastic differential geometry,” *Publications of the Research Institute for Mathematical Sciences*, vol. 17, no. 1, pp. 235–244, 1981.
- [8] E. Jørgensen, “Construction of the brownian motion and the ornstein-uhlenbeck process in a riemannian manifold on basis of the gangolli-mc.kean injection scheme,” *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 44, no. 1, pp. 71–87, 1978.
- [9] R. M. Dowell, *Differentiable approximations to Brownian motion on manifolds*. PhD thesis, University of Warwick, 1980.
- [10] X.-M. Li, “Random Perturbation to the Geodesic Equation,” *Ann. Prob.*, vol. 44, no. 1, pp. 544–566, 2016.
- [11] J. Angst, I. Bailleul, and C. Tardif, “Kinetic brownian motion on riemannian manifolds,” *arXiv preprint arXiv:1501.03679*, 2015.
- [12] J.-M. Bismut, “The hypoelliptic laplacian on the cotangent bundle,” *Journal of the American Mathematical Society*, vol. 18, no. 2, pp. 379–476, 2005.
- [13] J.-M. Bismut, “Hypoelliptic laplacian and probability,” *J. Math. Soc. Japan*, vol. 67, pp. 1317–1357, 10 2015.
- [14] E. Nelson, *Dynamical Theories of Brownian Motion*. Mathematical Notes - Princeton University Press, Princeton University Press, 1967.

- [15] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization*. Texts in Applied Mathematics, Springer New York, 2008.
- [16] P. Hänggi, “Nonlinear fluctuations: The problem of deterministic limit and reconstruction of stochastic dynamics,” *Phys. Rev. A*, vol. 25, pp. 1130–1136, Feb 1982.
- [17] G. Volpe, L. Helden, T. Brettschneider, J. Wehr, and C. Bechinger, “Influence of noise on force measurements,” *Physical review letters*, vol. 104, no. 17, p. 170602, 2010.
- [18] J. M. Sancho, M. S. Miguel, and D. Dürr, “Adiabatic elimination for systems of brownian particles with nonconstant damping coefficients,” *Journal of Statistical Physics*, vol. 28, no. 2, pp. 291–305, 1982.
- [19] S. Hottovy, A. McDaniel, G. Volpe, and J. Wehr, “The Smoluchowski-Kramers Limit of Stochastic Differential Equations with Arbitrary State-Dependent Friction,” *Communications in Mathematical Physics*, vol. 336, no. 3, pp. 1259–1283, 2014.
- [20] D. P. Herzog, S. Hottovy, and G. Volpe, “The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction,” *Journal of Statistical Physics*, vol. 163, no. 3, pp. 659–673, 2016.
- [21] J. Birrell, S. Hottovy, G. Volpe, and J. Wehr, “Small Mass Limit of a Langevin Equation on a Manifold,” *ArXiv e-prints*, Apr. 2016.
- [22] R. Chetrite and K. Gawędzki, “Fluctuation relations for diffusion processes,” *Communications in Mathematical Physics*, vol. 282, no. 2, pp. 469–518, 2008.
- [23] K. Gawędzki, “Fluctuation relations in stochastic thermodynamics,” *arXiv preprint arXiv:1308.1518*, 2013.
- [24] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, Springer New York, 2014.
- [25] J. Ortega, *Matrix Theory: A Second Course*. University Series in Mathematics, Springer US, 2013.
- [26] R. M. Wilcox, “Exponential operators and parameter differentiation in quantum physics,” *Journal of Mathematical Physics*, vol. 8, no. 4, 1967.

- [27] R. Khasminskii, *Stochastic stability of differential equations*, vol. 66. Springer Science & Business Media, 2011.